
COURTNEY PAQUETTE, McGill University

Halting Time is Predictable for Large Models: A Universality Property and Average-case Analysis

Average-case analysis computes the complexity of an algorithm averaged over all possible inputs. Compared to worst-case analysis, it is more representative of the typical behavior of an algorithm, but remains largely unexplored in optimization. One difficulty is that the analysis can depend on the probability distribution of the inputs to the model. However, we show that almost all instances of high-dimensional data are indistinguishable to first-order algorithms. Particularly for a class of large-scale problems, which includes random least squares and one-hidden neural networks with random weights, the halting time is independent of the probability distribution. With this barrier for average-case analysis removed, we provide the first explicit average-case convergence rates showing a tighter complexity not captured by traditional worst-case analysis. Finally, numerical simulations suggest this universality property holds in greater generality. Joint work with Elliot Paquette, Fabian Pedregosa, and Bart van Merriënboer